

PATENT APPLICATION

**AUTOMATED ACQUISITION OF VIDEO TEXTURES ACQUIRED
FROM A DIGITAL CAMERA FOR MAPPING TO AUDIO-DRIVEN
DEFORMABLE OBJECTS**

Inventor(s):

ALAN PEEVERS, a citizen of the United States, residing at
1211 Delaware St.
Santa Cruz, CA 95060

ALAN SEEFELDT, a citizen of the United States, residing at
270 Highland Ave., Apt. 23
Somerville, MA 02143

Assignee:

CREATIVE TECHNOLOGY, LTD.
31 International Business Park
Creative Resource
Singapore 609221
Republic of Singapore

Entity: Large

TOWNSEND and TOWNSEND and CREW LLP
Two Embarcadero Center, 8th Floor
San Francisco, California 94111-3834
Tel: 415-576-0200

AUTOMATED ACQUISITION OF VIDEO TEXTURES ACQUIRED FROM A DIGITAL CAMERA FOR MAPPING TO AUDIO-DRIVEN DEFORMABLE OBJECTS

5

CROSS-REFERENCES TO RELATED APPLICATIONS

This application claims priority from provisional Application No. 60/165,097, filed November 12, 1999, the disclosure of which is incorporated herein by reference.

10

BACKGROUND OF THE INVENTION

The use of computers to play music has created a great interest in techniques of providing a visual accompaniment to the music, known as "eye candy." This interest has been accelerated by the internet and the availability of MP3 audio files which can be downloaded and played on MP3 players, such as WinAmp, Sonique, etc., which are also downloaded. Many of these players are configured to run plug-in visualization programs. Some of these players such as LAVA (Live Audio-Visual Animation), created by the assignee of the present invention, respond in real time to events based on characteristics of the music being played, e.g., transients (such as beats) or sounds in a selected frequency band (such as a high note played by flute).

These real time video musical accompaniments are generally presented in a hard-wired format with limited control by the viewer. The scenes are usually designed by professional programmers and downloaded in compiled format to the users. Typically an interface is provided with minimal controls such as switching between fixed scenes, varying preset camera angles, and toggling between preset object or scene attributes.

Typically, the art in the scene consists of shapes and mapped textures which are stored in the memory of the computer. A visualization product disclosed in a co-pending application entitled DATA STRUCTURE AND CONTROL INTERFACE FOR AN EVENT DRIVEN VISUAL PRESENTATION, filed November 12, 1999 (Application No. 60/165,059, which is hereby incorporated by reference for all purposes) describes a product allowing editing by a viewer to facilitate substitution of existing scene objects and textures with new objects and textures which may be downloaded from the internet.

Often the original scene is entertaining and interesting but the user has ideas on how to customize or vary the scene. Accordingly, developers of event driven music visualizations are continuing to refine their products.

5

SUMMARY OF THE INVENTION

According to one aspect of the present invention, a video signal from an external source is texture-mapped onto an object displayed in computer generated animation. The video signal may be generated in real time by a video camera coupled to the computer. Thus, for example, a user may map a video image of his/her face onto an object in the animation.

10

According to another aspect of the invention, a template is used to map features of the face onto corresponding regions of an object in the animation. This mapping facilitates altering the appearance of selected features of the texture-mapped video image of the face.

According to another aspect of the invention, the appearance of the texture-mapped video is altered when selected events are detected. These events may be detected in an audio signal being animated or in the video signal being texture-mapped.

According to another aspect of the invention, the event detected in the video signal may be the crossing over a threshold of the value of a parameter, for example, the luminosity, of the video signal.

According to another aspect of the invention, the video signal alterations include altering the luminosity of the entire signal or regions of the signal and altering the palette to change the appearance of the video signal.

According to another aspect of the invention, regions of the mapped video image can be elevated or depressed when a selected event is detected.

Additional features and advantages will be apparent in view of the following detailed description and appended drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

30

Fig. 1 is a block diagram of a preferred embodiment of the present invention; Fig. 2 is a diagram depicting a scene; Fig. 3 is a diagram depicting a preferred embodiment of a user interface; and Fig. 4 is a flow chart depicting the steps of deforming a mapped video in response to audio events.

DESCRIPTION OF THE SPECIFIC EMBODIMENTS

A preferred embodiment of the present invention will be described with reference to the LAVA audio-driven visualization generator manufactured by the assignee of the present application. However, the invention is not limited to LAVA but has broad applicability for improving event driven visualization in many contexts.

Sub A
DMS A20 15 Emb P7 20 Sub A27 Fig. 1 is a block diagram overview of a LAVA player utilizing an MV3 format. The LAVA player 10 is a compiled executable application which reads the MV3 file from memory. The MV3 file is described in the above-referenced patent application and is read by the LAVA player 10 to obtain parameters and other information needed to initialize, render, and animate a LAVA visualization.

In Fig. 1 a video camera 12 is coupled to the LAVA player 10. This camera generates a video or a single frame (still image) which can be mapped onto an object by the user.

Fig. 2 is a depiction of a typical LAVA scene. Several objects are rendered and images are mapped as textures on the objects utilizing 3D-graphics techniques well-known in the art. In LAVA the textures can be moved along the surface of an object in response to events in the audio signal being played. One type of deformation is a "crawl" where the texture is displaced on the surface of the object.

In the preferred embodiment, a video or frame (still image) from the digital camera 12 is mapped to an object in real time so that the image recorded by the digital camera is viewed on the object as a video texture. Thus, an image external to the computer will be mapped to an object in the 3D LAVA visualization.

Additional still images, generated by the digital camera, can also be mapped as textures on objects in the LAVA visualization. A special technique can be utilized for mapping a face onto an object. For certain shapes, such as an ellipsoid, the mapping of the face is more realistic if the axis of the face is aligned along the long axis of the ellipsoid. A user interface (UI) displays a template for aligning the face's orientation and scaling to match the target object to provide realistic texture mapping. The template also facilitates mapping deformations to features such as the eyes or nose.

Further, the template facilitates positioning various features of the image (e.g. eyes and mouth) so that they are centered on previously positioned deformations that have been defined on the surface of the target object.

For example, as depicted in Fig 3, the UI 28 includes an object selection window 30 displaying thumbnails representing objects in a current scene, a texture positioning window 34 and a camera button 36. When a specific object 38 is selected, an alignment template 39 specific for that object is placed in the texture positioning window 34.

- 5 Once the camera button 36 is pressed a video image from the camera will appear in the texture positioning window 34, with the alignment template 39 displayed as an overlay on top of it. The user will then move the camera's subject (for example, her head, if the camera is pointed at her) so that various features of the video image are coincident with elements of the alignment template 39. In the example of the head, the subject moves her head (or the
10 camera) until the three points in the alignment template 39 corresponding to two eyes and a nose line up with these same features in the video image. Once the desired alignment is achieved, a snapshot button 40 is pressed which then maps the video image to the object 38 with the correct alignment. Deformations defined on object 38 may then alter the eyes and nose specifically, since they are located at positions on the object's surface corresponding to the features in the alignment template 39.

A preferred embodiment utilizes several techniques, singly or in combination, to animate a scene that includes texture-mapped real time video. These techniques include:

1. deforming parts of the surface of the objects onto which selected regions of the video image are texture-mapped by moving the vertices of the corresponding part of the 3D object. For example, if the video image were a face then the selected regions could be the nose and the eye sockets defined relative to the template. The deformation could be defined to elevate the nose and recess the eye sockets. Alternatively, the selected regions could be defined in terms of average luminosity of groups of pixels. The deformation could be defined to elevate regions with a luminescence value higher than a threshold value or to recess areas with a luminescence value less than a threshold value. For example, a bright nose would be projected forward and dark eye sockets would recede into the head. The elevation or depression is accomplished moving the vertices of the 3D image;
- 25 2. modifying texture mapping coordinates to cause the texture-mapped video to crawl and/or twist across the surface of the object. A technique for altering texture mapping coordinates is described in a commonly, assigned co-pending patent application entitled **AUDIO DRIVEN TEXTURE AND COLOR DEFORMATIONS OF COMPUTER GENERATED GRAPHICS** (filed 1/26/00, Application No. 09/491,530, which is hereby incorporated by reference for all purposes); and

3. modifying image parameters of parts of the video image (e.g. contrast or hue) or changing the color palette to cause the image to appear black-and-white, show color contours, or appear “psychedelic.” Alternatively, parts of the image, e.g., the eyes or nose, defined relative to the template or by other means, could have contrast, luminosity, or color modified.

As depicted in Fig. 4, each of the above described deformations or modifications is initiated by a trigger signal generated from a detected event. These triggering signals can either be generated by events detected in the audio signal being animated, as disclosed in the commonly assigned, co-pending application entitled AUDIO DRIVEN VISUAL REPRESENTATIONS, (filed 8/12/99, Application No. 09/373,405, which is hereby incorporated by reference for all purposes) or events detected in the video signal. For example, an event could be defined as detecting a luminescence value above a defined threshold for any pixel in a video frame.

Other types of deformations could be similar to “transitions” used in advertising videos, e.g., morphing a plane image into a page turn or other morphings.

The invention has now been described with reference to the preferred embodiments. Alternatives and substitutions will now be apparent to persons of skill in the art. For example, the video or still frame may not be generated in real time but could be read from a storage medium such as a disk or CD. Accordingly, it is not intended to limit the invention except as provided by the appended claims.